Pergamon

0042-6989(95)00130-1

Vision Res. Vol. 36, No. 3, pp. 445–459, 1996

0042-6989/96 $15.00 + 0.00

A Geometric Basis for Measurement of Three-dimensional Eye Position Using Image Processing

STEVEN T. MOORE,*† THOMAS HASLWANTER,‡ IAN S. CURTHOYS,‡ STUART T. SMITH‡

Received 22 September 1994; in revised form 23 February 1995

Polar cross correlation is commonly used for determination of ocular torsion from video images, but breaks down at eccentric positions if the spherical geometry of the eye is not considered. We have extended this method to allow three-dimensional eye position measurement over a range of ±20 deg by determining the correct projection of the eye onto the image plane of the camera. We also determine the orientation of the camera with respect to the eye, allowing eye position to be represented in appropriate head-fixed coordinates. These algorithms have been validated using both *in vitro* and *in vivo* measures of eye position.

Eye movements Image processing Ocular torsion Video VOR

INTRODUCTION

For a full understanding of the consequences of eye movements it is necessary to measure horizontal, vertical and torsional eye movements. The preferred method for three-dimensional eye position measurement is still the scleral search coil technique (Robinson, 1963; Collewijn, Van der Steen, Ferman & Jansen, 1985), despite problems such as coil slippage (especially in torsion), limited testing time (about 40 min per session in our experience) and the use of expensive contact lenses. These limitations, and advances in digital image processing technology in recent years, have provided the stimulus for the development of video-based systems for measurement of three-dimensional eye position.

The small number of existing image processing systems for eye position measurement fall into two broad categories: those based on the tracking of natural or attached landmarks on the eye (Nakayama, 1974; Parker, Kenyon & Young, 1985; Yamanobe, Taira, Morizono, Yagi & Kamio, 1990; Ott, Gehle & Eckmiller, 1990), and systems employing some variation of the *polar cross correlation* method, where the iral intensity along a circular sampling path is used to determine ocular torsion (Hatamian & Anderson, 1983; Vieville & Masse, 1987; Clarke, Teiwes & Scherer, 1991; Moore, Curthoys & McCoy, 1991). A mathematical basis for landmark tracking systems has been developed by Nakayama (1974), representing eye position in Fick coordinates. A system based on this method has been realized using a contact lens to provide the markers (Ott *et al.*, 1990), but has similar disadvantages to scleral search coils. Landmark tracking systems which rely on natural features of the eye face the difficulty that such features may not be prominent in all subjects.

The polar cross correlation method relies on the naturally occurring radial pattern of the iris and is arguably more applicable to a wider range of subjects, though variation in the strength of radial iral patterns does exist between individuals. Polar cross correlation accurately measures ocular torsion in the absence of horizontal and vertical deviations of the eye, but breaks down at eccentric eye positions if the spherical geometry of the eye is not considered. We have extended the polar cross correlation method by developing an algorithm to compensate for the geometric distortion of the projection of the eye onto the *image plane* of the camera. We have also developed analytical formulae and calibration procedures to determine the spatial orientation of the eye with respect to the camera, which allow representation of eye position in appropriate head-fixed coordinates regardless of camera alignment. This has enabled accurate measurement of three-dimensional eye position using polar cross correlation over a large range of eye movements. We have validated this system by comparing results for geometrically compensated and uncompensated torsion measures from an artificial eye with a numerical simulation, and by direct measures on human

*Department of Electrical Engineering, University of Sydney, Sydney, NSW 2006, Australia.

†To whom all correspondence should be addressed at: Eye and Ear Research Unit, Department of Neuro-otology, Royal Prince Alfred Hospital, Missenden Road, Camperdown NSW 2050, Australia [*Email* stevenm@psychvax.psych.su.oz.au].

‡Department of Psychology, University of Sydney, Sydney, NSW 2006, Australia.

subjects. The errors induced by using an orthographic rather than a central (perspective) projection to model the imaging process of the camera are also considered, as well as the effect of the cornea on the projection of the eye onto the image plane.

THEORETICAL BACKGROUND

The polar cross correlation method

The polar cross correlation method relies on the fact that most of the variation in pixel intensity of a digitized image of the iris occurs in the angular direction in a polar coordinate system centred on the pupil (Hatamian & Anderson, 1983). Measurement of ocular torsion can therefore be reduced to a one-dimensional signal processing task by forming an *iral signature* $I(\alpha)$ from the pixel intensity values of the iris along a circular *sampling path* centred on the pupil, where α is the polar angle of a point on this circle. In practice the sampling of the iris is often limited to a segment of this circle, as upper eyelid droop and tear build up in the lower eye render these areas unsuitable for obtaining an iral signature.

The sampling path is 1 pixel wide and is maintained at approximately the same angular position on the iris by tracking the pupil centre and translating the sampling path accordingly. Horizontal and vertical eye position is determined from the orientation of the optical axis of the eye, which is approximately given by the line passing through the centre of the eye and the pupil centre. In previous implementations of the polar cross correlation method the assumption has been made that the shape of the sampling path remains constant at all eye positions (Hatamian & Anderson, 1983; Vieville & Masse, 1987; Clarke *et al.*, 1991; Moore *et al.*, 1991). Iral signatures from each video frame are cross correlated with an *iral reference signature* $I_{ref}(\alpha)$, obtained at the start of testing. In practice, the cross correlation function is often calculated using fast Fourier transforms (FFTs),

$$\mathrm{Corr}(I(\alpha), I_{ref}(\alpha)) = F^{-1}\{F\{I(\alpha)\} \cdot F\{I_{ref}(\alpha)\}^*\} \quad (1)$$

where the asterisk indicates the complex conjugate. The shift in the peak of this cross correlation function indicates the torsional position of the eye for each video frame relative to the reference image.

As the results presented below demonstrate, iral signatures obtained using purely translated (geometrically *uncompensated*) sampling paths are distorted and produce erroneous torsion results. To correctly measure torsion at eccentric eye positions, it is necessary to take the horizontal and vertical orientation and the geometry of the eye into consideration when forming the sampling path in the image plane. In the following sections we describe a geometric basis for accurate three-dimensional eye position measurement from video images, which determines the correct (geometrically *compensated*) sampling path for any horizontal and vertical eye position.

Imaging geometry

In our analysis the eye is assumed to be a perfect sphere exhibiting ideal *ball and socket* behaviour. All eye movements are therefore pure rotations around the centre of this sphere, with no translational component. The validity of this assumption will be considered in the discussion. The iris is modelled as a plane section (Wolff, 1940), centred on the pupil, at a distance r_p from the centre of the eye. In the following derivation, matrices are represented by outline uppercase characters (e.g. $\mathbb{R}$), points in three-dimensional space by underscored uppercase characters (e.g. $\underline{P}$), and unsigned scalar quantities by lowercase italicized characters (e.g. r_p, f, d).

To derive the mathematical formulae for the projection of the eye onto the image plane we define an orthogonal head-fixed, right-handed coordinate system $\{\underline{h}_1, \underline{h}_2, \underline{h}_3\}$, with the origin at the centre of the eye as shown in Fig. 1. The $\underline{h}_2$–$\underline{h}_3$ plane is parallel to the coronal plane, with the $\underline{h}_2$ axis parallel to the interaural axis of the subject. The $\underline{h}_1$ axis is perpendicular to $\underline{h}_2$ and

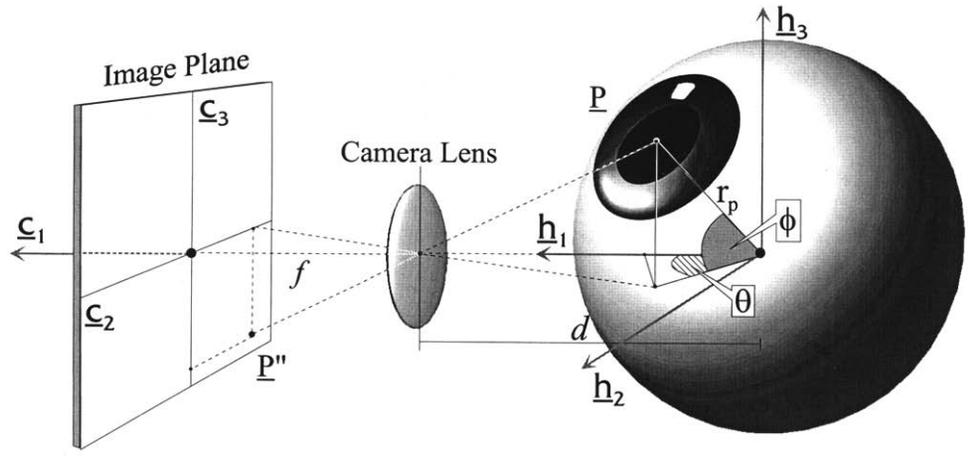

FIGURE 1. Central projection of the eye onto the image plane. $\{\underline{h}_1, \underline{h}_2, \underline{h}_3\}$ and $\{\underline{c}_1, \underline{c}_2, \underline{c}_3\}$ are the basis of the head- and camera-fixed coordinate systems, with the head-fixed frame translated by $\underline{T} = (-(f+d), 0, 0)$ with respect to the camera-fixed system. $\underline{P}$ is the pupil centre in the head-fixed frame, $\underline{P}''$ is the projection onto the image plane and (θ, ϕ) are the horizontal and vertical Fick components of eye position in head-fixed coordinates.

$\underline{h}_3$. The *reference position* of the eye is defined as the position where the centre of the pupil lies on the $\underline{h}_1$ axis.

We also define a camera-fixed coordinate frame $\{\underline{c}_1, \underline{c}_2, \underline{c}_3\}$, where the $\underline{c}_2$ and $\underline{c}_3$ axes lie within the image plane, and $\underline{c}_1$ corresponds to the "line of sight" of the camera. The plane of the camera lens is located on the $\underline{c}_1$ axis at a distance f from the image plane, and a distance d from the centre of the eye respectively. When the camera is focused on distant objects, f is equal to the focal length of the lens. The coordinates of a point in space in the head-fixed frame, e.g. the pupil centre $\underline{P} = (p_1, p_2, p_3)$, and the corresponding coordinates with respect to the camera-fixed system, $\underline{P}' = (p_1', p_2', p_3')$, are related by

$$\underline{P}' = \mathbb{R}_{off} \cdot \underline{P} + \underline{T} \tag{2}$$

where $\mathbb{R}_{off}$ and $\underline{T}$ are the rotation and translation of the head-fixed coordinate frame with respect to the camera coordinate system. Due to the non-commutativity of finite rotations it is necessary to specify a rotation convention when decomposing $\mathbb{R}_{off}$ into a horizontal, vertical and a torsional rotation. We have adopted the Fick sequence (Fick, 1854), which is commonly used in oculomotor research. A description of this rotation sequence and the expanded form of rotation matrices are given in Appendix A.

The central (perspective) projection of $\underline{P}'$ onto the image plane, $\underline{P}''$, is given by

$$\underline{P}'' = \begin{pmatrix} 0 \\ x \\ y \end{pmatrix} = f * \begin{pmatrix} 0 \\ \dfrac{p_2'}{f + p_1'} \\ \dfrac{p_3'}{f + p_1'} \end{pmatrix} \tag{3}$$

where x and y are the components of $\underline{P}''$ along the $\underline{c}_2$ and $\underline{c}_3$ axes respectively. Note that the definition of $\{\underline{c}_1, \underline{c}_2, \underline{c}_3\}$ means that p_1' is always negative for points in front of the camera lens. This projection is shown in Fig. 1. An extensive description of imaging geometry can be found in Gonzalez and Wintz (1987).

Video system calibration

In order to measure three-dimensional eye position from video images, the system must first be calibrated to determine $\mathbb{R}_{off}$ and $\underline{T}$, the rotation and translation of the head-fixed coordinate frame with respect to the camera. The distance r_p from the pupil centre to the centre of the eye must also be calculated. This can be accomplished as follows.

(i) The subject is asked to fixate on a number of points at known horizontal and vertical Fick angles in head-fixed coordinates (θ, ϕ). The combined horizontal and vertical rotation of the eye from the reference position to the current position is described by $\mathbb{R}_{\theta\phi}$, given in Appendix A. Note that the torsional position of the eye does not affect the calibration process.

(ii) At each calibration position, the centre of the pupil in camera-fixed coordinates, $\underline{P}'$, is given by

$$\underline{P}' = \mathbb{R}_{off} \cdot \mathbb{R}_{\theta\phi} \cdot \underline{P}_0 + \underline{T} \tag{4}$$

where $\underline{P}_0 = (r_p, 0, 0)$ is the position of the pupil centre in head-fixed coordinates when the eye is in the reference position.

(iii) The centre of the pupil in the image plane $\underline{P}'' = (0, x, y)$ is calculated for each calibration position. $\underline{P}''$ can be determined using a number of methods: averaging the coordinates of the edge points of the pupil (Hatamian & Anderson, 1983); fitting a circle to the pupil boundary (Sung & Anderson, 1991); calculating the "*centre of mass*" of the pupil (where each pixel within the pupil boundary is assigned a mass of 1). The first two techniques assume a circular form for the pupil, but in our experience the pupil can deviate markedly from this shape. The centre of mass algorithm determines the centre of any arbitrary shape, and is used in our implementation.

(iv) Using the coordinates of these pupil centres in the image plane and equations (3) and (4), $\mathbb{R}_{off}$, $\underline{T}$ and r_p can be determined. In general this will require numerical techniques. If the distance d between the centre of the eye and the plane of the camera lens is much larger than r_p, an orthographic projection can be used instead of the more complex perspective projection (see Appendix B), and an analytical solution can be derived (Appendix C).

(v) For the measurement of ocular torsion, it is necessary to acquire an iral reference signature when the eye is in the reference position. This is obtained by sampling along a circle which is located on the iral plane and centred on the pupil with a radius r_{ip} (as shown in Fig. 2). In the head-fixed coordinate system, a point S_0 on this circle is represented in

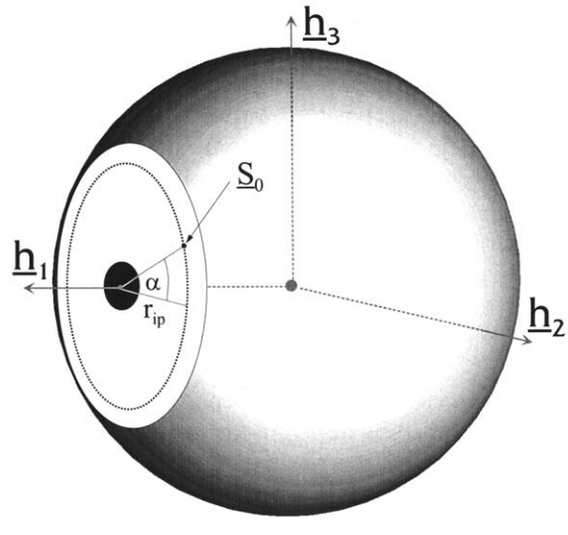

FIGURE 2. Diagram of the eye in the reference position. The iris is modelled as a plane section, centred on the pupil. $\underline{S}_0$ is a point on the sampling circle, which is represented in parametric form as a function of the angle α about the $\underline{h}_1$ axis, and is at a radius r_{ip} from the pupil centre.

parametric form as a function of the angle α about the $\underline{h}_1$ axis, and is given by

$$\underline{S}_0 = \begin{pmatrix} r_p \\ r_{ip} * \cos(\alpha) \\ r_{ip} * \sin(\alpha) \end{pmatrix}. \quad (5)$$

The location of this circle in camera-fixed coordinates, $\underline{S}'$, is obtained by substituting $\underline{S}_0$ into equation (2). The location of the geometrically compensated sampling path $\underline{S}''$ in the image plane is given by applying the central projection of equation (3) to $\underline{S}'$, and the iral reference signature $I_{ref}(\alpha)$ is formed by sampling along $\underline{S}''$. In practice a *sampling arc* is often used instead of the entire circle by choosing the appropriate α values for the initial and end points of the arc.

Determining three-dimensional eye position

After the video system has been calibrated, it is possible to determine three-dimensional eye position (in Fick angles) from subsequent video frames as follows.

(i) The centre of the pupil in the image plane, $\underline{P}''$, is determined using the centre of mass algorithm.

(ii) If f, r_p and $\underline{T}$ are known, $\underline{P}'$ (the pupil centre in camera-fixed coordinates) can be derived from $\underline{P}''$ (as shown in Appendix D).

(iii) The horizontal and vertical Fick angles of the current eye position in the head-fixed coordinate system, (θ, ϕ), can be calculated by inverting equation (4)

$$\underline{P} = \mathbb{R}_{\theta\phi} \cdot \underline{P}_0 = r_p * \begin{pmatrix} \cos(\theta)\cos(\phi) \\ \sin(\theta)\cos(\phi) \\ -\sin(\phi) \end{pmatrix}$$

$$= \mathbb{R}_{off}^{-1} \cdot (\underline{P}' - \underline{T}). \quad (6)$$

(iv) To accurately measure ocular torsion it is necessary to determine the correct projection of the sampling arc onto the image plane (i.e. the sampling path) at the current eye position, rather than simply translating the original sampling path in the image plane by the same amount as the pupil centre, as shown in Fig. 3. The position of the sampling arc in the camera-fixed coordinate system at this eye position, $\underline{S}'$, is given by

$$\underline{S}' = \mathbb{R}_{off} \cdot \mathbb{R}_{\theta\phi} \cdot \underline{S}_0 + \underline{T} \quad (7)$$

and the geometrically compensated sampling path $\underline{S}''$ is obtained by applying the central projection of equation (3).

(v) The iral signature $I(\alpha)$ is obtained by sampling along $\underline{S}''$. The cross correlation of $I(\alpha)$ with $I_{ref}(\alpha)$ is calculated using equation (1), and the shift in the peak of this function provides the Fick torsion ψ of the eye. A practical implementation of equations (6) and (7) is discussed in Appendix D.

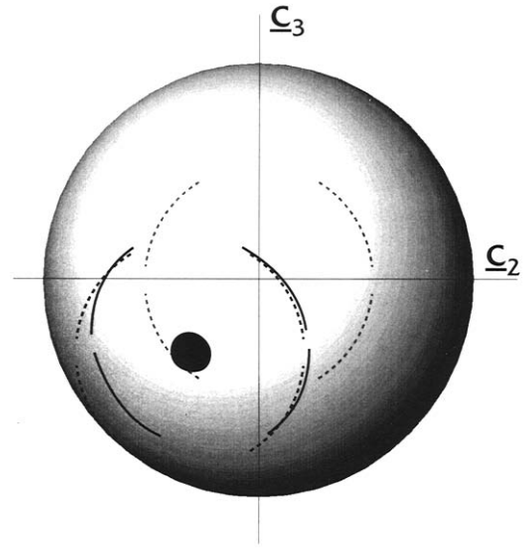

FIGURE 3. Diagram of the effect of geometric distortion of the sampling path, with the line of sight directed 20° left and 20° up. Note that the image is inverted due to the central projection. The light dashed lines indicate the original location of four sampling paths (one in each quadrant) when the eye is in the reference position. Dark dashed and solid lines indicate the geometrically uncompensated (translated) and compensated sampling paths respectively. The deformation of the uncompensated sampling path depends on eye position, as well as on the original location of the sampling arc $\underline{S}_0$. In practice the sampling paths from the lower two quadrants would be occluded by the bottom eyelid, and iral signatures would have to be obtained from the upper quadrants.

Simulation of the effect of rotational offset

If the effect of the rotation $\mathbb{R}_{off}$ of the head-fixed coordinate system with respect to the camera-fixed frame is ignored, errors will be induced in measured eye position. To investigate this effect we have developed a simulation of the projection of the pupil centre onto the image plane for various rotational offsets using a camera set-up typical of the Video Torsion Measurement system (Moore *et al.*, 1991). The centre of the eye is translated by $\underline{T} = (-(f + d), 0, 0)$ with respect to the camera-fixed coordinate system, where $f = 12.5$ mm, $d = 72$ mm and the radius of the eye at the pupil centre r_p is 12 mm (Wolff, 1940). The results of this simulation are shown in Fig. 4.

Using equations (3) and (4), we have simulated horizontal, vertical and torsional rotational offsets $(\theta_{off}, \phi_{off}, \psi_{off})$ of $(5°, 0°, 0°)$, $(0°, 5°, 0°)$, $(5°, 5°, 0°)$ and $(0°, 0°, 5°)$. In each case the central projection of the pupil centre onto the image plane was determined over a range of horizontal and vertical eye positions ($\pm 20°$ in θ and ϕ at 5° intervals). The errors induced in determining the horizontal and vertical Fick angles of eye position from these shifted pupil centres were then calculated using equation (6), with $\mathbb{R}_{off} = \mathbb{I}$ (i.e. the unit matrix). For a purely horizontal offset angle of 5° [Fig. 4(A)], the error magnitude (mean and SD) in horizontal eye position was $\theta_{error} = 0.38 \pm 0.26°$ (maximum 1.07°), and in vertical eye position $\phi_{error} = 0.04 \pm 0.04°$ (maximum 0.15°); for a purely vertical offset of 5° [Fig. 4(B)], $\theta_{error} = 0.07 \pm 0.07°$ (maximum 0.29°) and

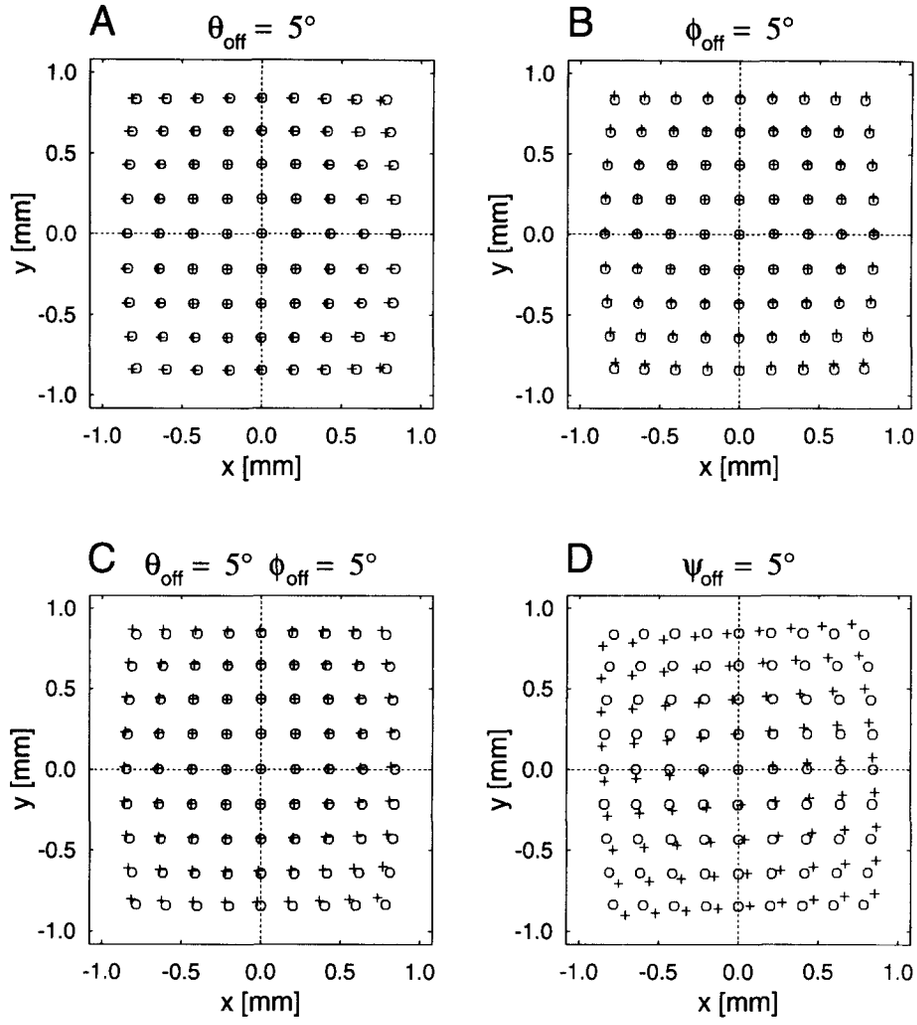

FIGURE 4. Simulation of the effect of rotational offset of the head-fixed coordinate system with respect to the camera on the central projection of the pupil centre $\underline{P}'' = (0, x, y)$ onto the image plane. The location of the pupil centre has been calculated for horizontal and vertical positions between $\pm 20°$ in 5° steps. ○ The pupil centre coordinates with no rotational offset; × pupil centre location with a rotational offset (θ_{off}, ϕ_{off}, ψ_{off}) of (A) (5°, 0°, 0°), (B) (0°, 5°, 0°), (C) (5°, 5°, 0°) and (D) (0°, 0°, 5°). Note that the image plane axes have been inverted to reflect eye movements in space.

$\phi_{error} = 0.36 \pm 0.24°$ (maximum 1.01°); for a combined horizontal and vertical offset of 5° [Fig. 4(C)], $\theta_{error} = 0.38 \pm 0.30°$ (maximum 1.54°) and $\phi_{error} = 0.36 \pm 0.26°$ (maximum 1.19°). Note that the maximum error in all the above cases occurred at highly eccentric eye positions when the eye was looking in the same direction as the offset, i.e. to the left [Fig. 4(A)], down [Fig. 4(B)] and to the left and down [Fig. 4(C)]. The error was significantly less in all other directions. Large errors were induced over the entire range of the eye positions for a purely torsional camera offset of 5° [Fig. 4(D)], with $\theta_{error} = 0.97 \pm 0.58°$ (maximum 1.84°) and $\phi_{error} = 0.96° \pm 0.56°$ (maximum 1.78°).

These results demonstrate that the largest errors in the determination of eye position are induced by a torsional camera offset, and that for horizontal or vertical offsets the error is significantly less. In practice horizontal and vertical positioning of the camera is readily achieved within the 5° offset values simulated above. In this case it may be possible to ignore the horizontal and vertical camera offsets, but the effect of a torsional offset about the optical axis of the eye must be taken into account. A practical implementation of this simplified calibration procedure is described in the Experimental Methods section.

Corneal optics

Light rays coming from the iris onto the image plane must pass through the cornea and the camera lens. Using Gullstrand's schematic eye as a model (Gullstrand, 1909), the cornea is around 12 mm in diameter and the anterior and posterior surfaces are spherical to a first approximation with radii of 7.7 and 6.8 mm respectively. The cornea may be approximated by a single thin lens, because a pair of thin lenses can be determined that have the same optical effect as the anterior and posterior refractive surfaces, and the positions of these equivalent lenses coincide within 0.001 mm. The optics of the cornea can therefore be modelled as a thin 43-D lens positioned approx. 0.5 mm in front of the corneal vertex. This gives rise to a slightly magnified image of the iris and pupil (around 13% larger) which appears to float

approx. 1 mm in front of and parallel to the actual iris (Bennett & Rabbetts, 1984). For a video-based measurement system, this optical effect would be reflected in a correspondingly larger value of r_p, the distance from the centre of the eye to the pupil centre.

EXPERIMENTAL METHODS

The video torsion measurement system

To accurately measure three-dimensional eye position using polar cross correlation over a large range of horizontal and vertical eye positions ($\pm 20°$), we have implemented the algorithms described above in a PC-based image processing system called *video torsion measurement* (VTM) (Moore *et al.*, 1991). This system was initially designed for accurate measurement of ocular torsion within a small range of horizontal and vertical eye positions (approx. $\pm 5°$).

Approximately 30 min prior to testing 2–4 drops of the pupil constrictor 2% pilocarpine hydrochloride (Isoptocarpine™, Alcon Laboratories) are administered to the subject's eye, ensuring that there is no movement of the iris during testing due to fluctuating pupil size. Variations in pupil size would lead to changes in the striation pattern of the iris, which in turn cause changes in the iral signature unrelated to ocular torsion, inducing errors in the torsional component of measured eye position. The use of a pupil constrictor is not a requirement for video-based three-dimensional eye position measurement, but enables ocular torsion to be measured with a higher degree of accuracy. Visual acuity is reduced, and if complex visual stimuli are presented rather than the primarily vestibular stimuli used in our research, the pupil would have to be unconstricted.

An image matrix composed of 512 × 512 8-bit pixel intensity values is formed by digitizing a standard video frame using an image processing card (Matrox MVP-AT) placed in a PC. The video signal is provided by a small ("lipstick") infra-red sensitive CCD video camera (Panasonic WV-CD1E) mounted vertically on custom built goggles worn by the subject. The eye is illuminated by two bundles of six infra-red LEDs (CQY89A), and an infra-red sensitive mirror (OCLI, Santa Rosa, Coolbeam mirror), which is transparent in the visible range, reflects the image of the eye onto the CCD of the camera. This allows the subject a wide field of view during recording. The VTM camera set-up is shown in Fig. 5.

To facilitate the real-time implementation of the three-dimensional eye position measurement algorithm described above, we have used an orthographic projection in place of the central projection of equation (3) to model the camera imaging process. The errors induced by using this simplification are discussed in Appendix B.

Prior to testing the subject is asked to look at a fixation point (the *reference LED*) positioned directly ahead. This eye orientation is taken as the reference position for all subsequent calculations. The pupil is identified using a threshold function (all pixels darker than a certain intensity level are designated as belonging

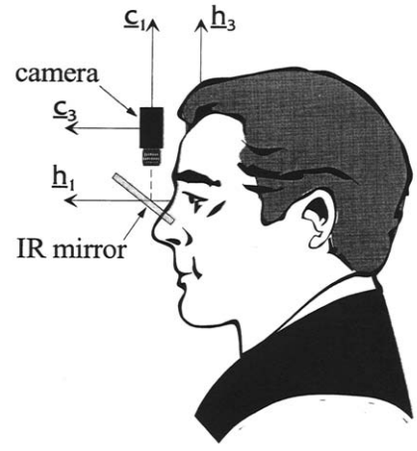

FIGURE 5. Diagram of the imaging set-up for the VTM system. The video camera is mounted vertically, and the image of the eye is reflected off an infra-red mirror.

to the pupil) and the centre of the pupil is calculated using the centre of mass algorithm. The VTM system is then calibrated to determine r_p, the radius of the eye at the pupil centre, and the orientation of the eye relative to the camera. The subject is asked to fixate on horizontally aligned calibration LEDs at $\pm 10°$ and $\pm 20°$. At each fixation point the system acquires 20 images and the mean and SD of the location of the pupil centre in the image plane is calculated. This reduces the effect of random biological "jitter" and provides a more accurate estimate of the pupil centre than a single image. The SD of the values for the pupil centre can be used to reject an estimate for a particular fixation point if a set error threshold is exceeded.

From the location of the pupil centre in the image plane at each calibration angle, r_p is calculated using equation (6). The torsional offset ψ_{off} of the head-fixed coordinate system with respect to the camera is derived from the gradient of the line of best fit through the location of the pupil centre in the image plane at the four calibration angles and the reference position. The effects of horizontal and vertical offsets are minimized through precise positioning of the camera, and are ignored (i.e. $\theta_{\text{off}} = \phi_{\text{off}} = 0°$). The effect of the torsional offset on the horizontal and vertical components of eye position can be removed after testing (see Appendix D). Note that the torsional component of eye position is measured relative to an iral reference signature using the polar cross correlation method, and is independent of the torsional offset of the camera.

The projection of the pupil centre with the eye in the reference position forms the origin of the image plane coordinate system. The image is divided into quadrants designated as *upper left*, *lower left*, *lower right* and *upper right*, as seen from the subject's point of view. With the eye in the reference position, a sampling path is formed in each quadrant, corresponding to a sampling arc on the iris at a radius r_{ip} from the pupil centre, as shown in Fig. 6. Each sampling path is 256 pixels long and typically spans around 70°. The iral intensity

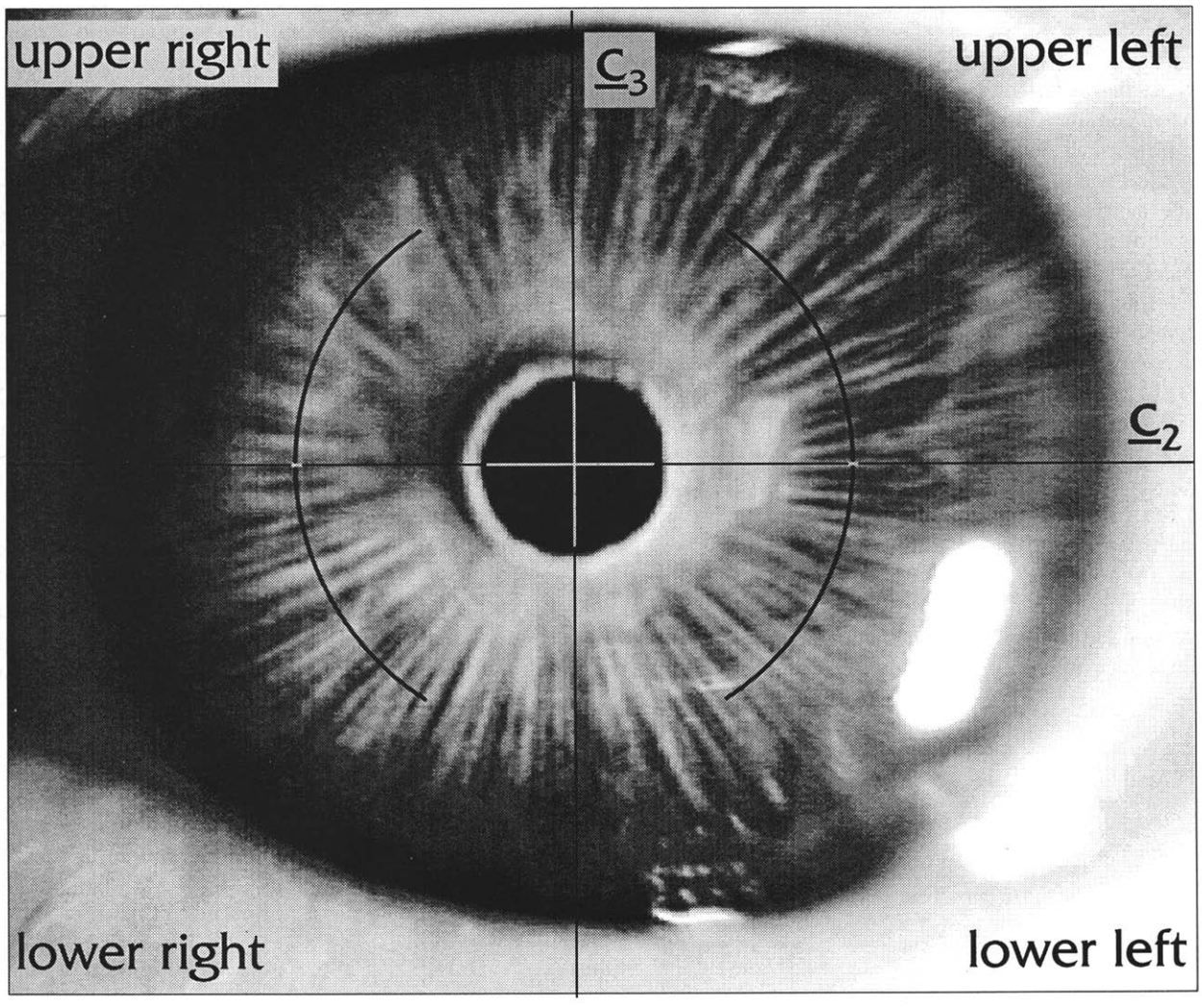

FIGURE 6. Video image from the VTM system. The eye is in the reference position, and the location of the four sampling paths (one in each quadrant) are shown.

values along each sampling path form the iral reference signature for each quadrant.

The torsional resolution of the system is primarily determined by the radius of the sampling path (in pixels) r_{sp}, and in our system is given to a first approximation by the empirical relationship

$$\psi_{res} = 0.4 - \frac{r_{sp}}{600}. \quad (8)$$

Equation (8) expresses the torsional resolution in degrees, which usually ranges from 0.1° to 0.2°. The horizontal and vertical resolution is dependent on the size of the eye in the video image, and for a typical VTM camera set-up the resolution is 0.15°.

During testing horizontal and vertical eye position are determined as described above, and iral signatures are taken from a geometrically compensated sampling path formed in the quadrant diametrically opposed to the current pupil location. These values are stored for off-line processing. This involves interpolation using a cubic spline function and re-sampling of the iral signatures to provide 512 data points with equally spaced angular increments. A high pass filter removes low frequency lighting artifacts. Iral signatures obtained during testing are then cross correlated with the iral reference signature from the appropriate quadrant, and a least squares quadratic fit is performed about the maximum of the cross correlation function to increase resolution. The shift in the peak of this function is a measure of the amount of ocular torsion. Eye position is represented in Fick coordinates (three angles denoting horizontal, vertical and torsional position) but these are easily converted into other representations such as Helmholtz coordinates (Helmholtz, 1867), quaternions (Westheimer, 1957) and rotation vectors (Haustein, 1989).

Artificial eye and fick gimbal

To test the accuracy of the VTM system and the geometric compensation algorithm we have constructed a calibration device consisting of an artificial eye fixed to a Fick gimbal. The artificial eye consisted of a white plastic hemisphere having dimensions similar to that of a human eyeball (radius 15 mm). A 3 mm diameter hole

was drilled in the centre of this hemisphere to act as a "pupil", and a sheet of black heat shrink tubing placed behind the hole prevented light from shining through. A number of radial lines scribed around the pupil simulated a human iral striation pattern. The artificial eye was positioned such that the centre of the hemisphere coincided with the centre of the Fick gimbal. Rotation of the artificial eye was possible about the horizontal and vertical axes of the gimbal with an accuracy of 0.25°. Torsional rotation around the "line of sight" of the artificial eye was not possible with this set-up, although experimental results using a precision torsion device have been published previously (Moore *et al.*, 1991).

The gimbal and video camera were mounted on an optical bench. They were aligned such that the line of sight was orthogonal to the image plane of the camera, and the pupil was approximately centred in the image plane when the artificial eye was in the reference position. The torsional position of the camera was adjusted so that horizontal or vertical rotations of the gimbal corresponded to horizontal or vertical movements of the artificial eye in the image plane. The artificial eye was illuminated with two bundles of LEDs as described in the VTM methods section. The gimbal and camera mounts could be translated in all three dimensions to obtain a centred, focused image.

Scleral search coil system and human centrifuge

To verify results for torsional eye position obtained from the VTM system, we have simultaneously recorded eye position data with the well established dual search coils technique (Robinson, 1963; Collewijn *et al.*, 1985). The magnetic fields were driven by a coil system based on a non-resonant circuit designed by Remmel (1984). An algorithm developed by Merfeld and Young (1992) was used to calculate three-dimensional eye position from the recorded voltages and expressed as rotation vectors.

To elicit ocular torsion, the subject was seated on a fixed chair human centrifuge (Servo-Med CF10) 1 m from the centre of rotation, left ear facing out. With head and body firmly restrained, the subject was accelerated counterclockwise (CCW) to a constant velocity. Testing was carried out in darkness except for

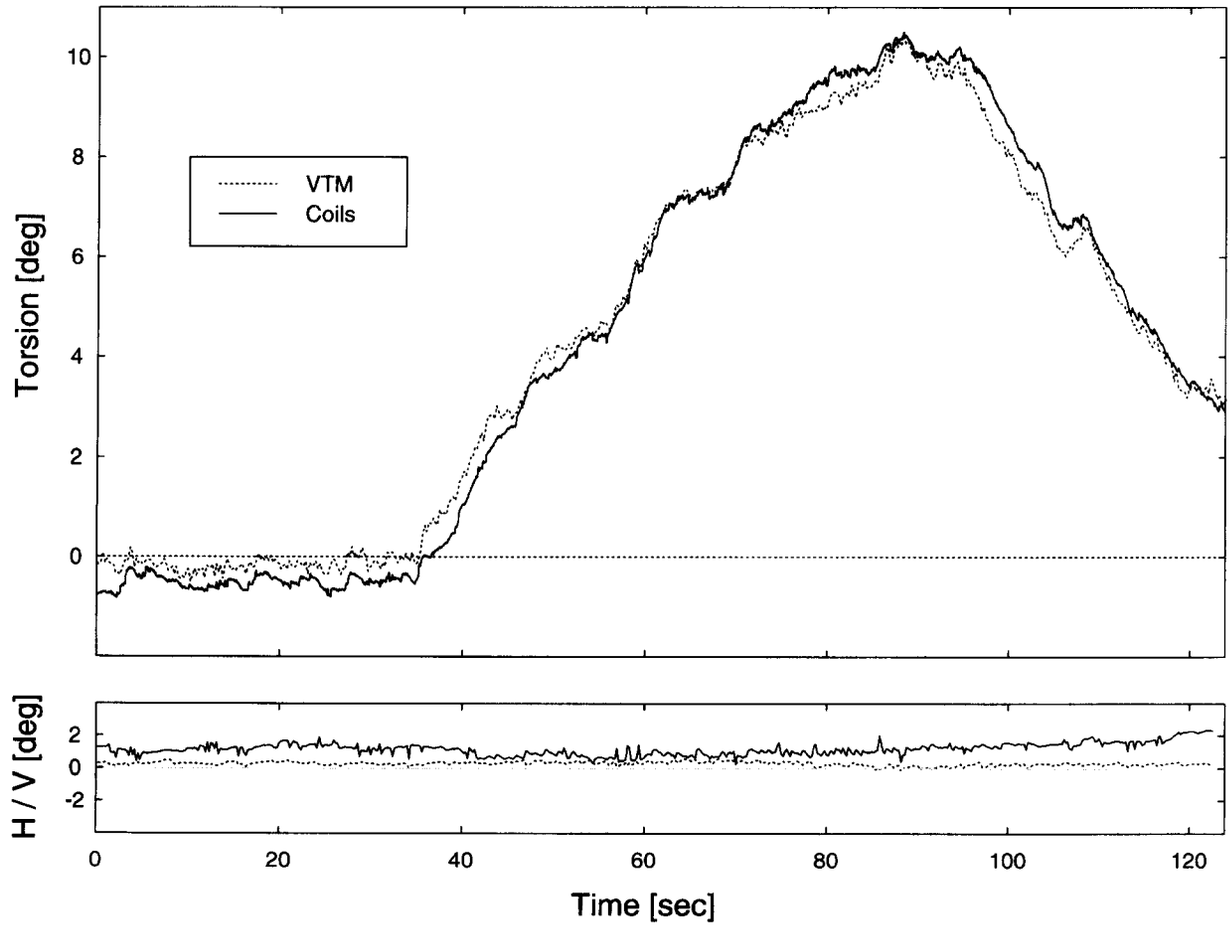

FIGURE 7. Simultaneous video/scleral search coil recordings of eye position for a human subject undergoing acceleration on the arm of a centrifuge. The upper plot shows the ocular torsion measured by the VTM system (dashed line) and coils (solid line), synchronized by aligning the maximum value for both recordings. The lower trace shows the rotation vector components for horizontal (solid line) and vertical (dashed line) eye position expressed in degrees, determined from the coil voltages. Horizontal and vertical eye movements were suppressed by a fixation LED as the coil annulus interfered with infrared illumination of the iris at eccentric eye positions.

a reference LED positioned 60 cm in front of the subject.

RESULTS

Simultaneous video/scleral search coil torsion recordings

We have performed an *in vivo* test of video-based measurement of ocular torsion by simultaneous VTM and scleral search coil recordings of eye position from a subject undergoing rotation on the arm of a centrifuge. Following calibration of the VTM and coil systems, simultaneous VTM/coils recordings of the position of the left eye were carried out. The subject was stationary for the first 30 sec, then accelerated CCW (back to motion) at 2.5°/sec² from rest to 150°/sec. After 30 sec at constant velocity eye position recording ceased.

The reference LED was turned on throughout the experiment to facilitate the suppression of horizontal and vertical eye movements, as at eccentric eye positions the coil annulus interfered with the illumination of the iris. The torsional components of eye position measured by VTM and the coils system are shown in Fig. 7, and the results for both systems are in close agreement.

Pupil-centre determination

The centre of mass algorithm for determining the pupil centre was validated *in vitro* using the artificial eye described above, placed in horizontal and vertical positions over a range of ±20° in 5° steps. At each orientation 10 images were acquired using the VTM system, and the horizontal and vertical Fick components of the eye position calculated from the pupil centre coordinates. The results of this experiment are shown in Fig. 8. The error magnitudes were small: mean and SD of the errors for all 810 pupil centre coordinates were 0.12 ± 0.09° for the horizontal component (maximum error 0.48°), and 0.16 ± 0.10° for the vertical component (maximum error 0.44°).

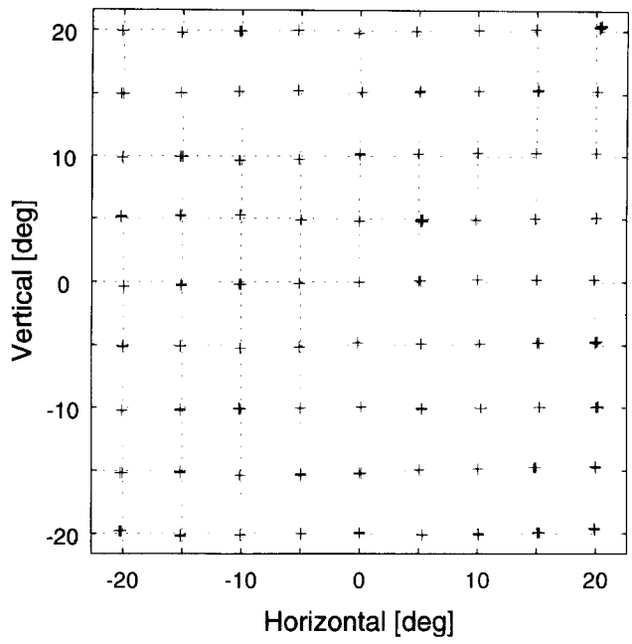

FIGURE 8. *In vitro* validation of the centre of mass algorithm for determining the pupil centre. Horizontal and vertical Fick components of the position of an artificial eye mounted on a Fick gimbal were determined with the VTM system. Ten pupil centre coordinates were acquired at each eye orientation, and measured values closely correspond to actual eye positions.

Validation of the geometric compensation algorithm

To verify the geometric compensation algorithm described by equations (5) and (7), VTM measures of torsion position were obtained *in vitro* from the artificial eye described above. The eye was placed in horizontal and vertical positions over a range of ±20° in 5° steps. The torsional resolution of the VTM system was 0.14°. At each eye position iral signatures were simultaneously sampled from a geometrically compensated and uncompensated sampling path (see Fig. 3) from the quadrant diametrically opposed to the current location of the pupil centre in the image plane. For all orientations the eye was placed in a position with 0° Fick torsion, and any torsion indicated by the VTM system was in error.

The measured torsion values were compared to values obtained from a numerical simulation of the polar cross correlation method (Haslwanter & Moore, 1995). The simulation was based on the assumption of a purely radial iral pattern (Hatamian & Anderson, 1983). An iral pattern was generated using the main frequency components of an iral signature from the VTM system, sampled from the artificial eye used in this experiment. The projection of the eye onto the image plane was simulated over a range of horizontal and vertical eye positions, and at each orientation iral signatures from compensated and uncompensated sampling paths were calculated using this simulated iral pattern. These signatures were then cross correlated with the iral reference signature to obtain the simulated torsion errors. The angular position and radius of the simulated sampling paths with respect to the pupil centre were identical to those used by the VTM system.

The results of the experimental and simulated torsion errors are shown in Fig. 9. Each torsion value for the experimental plots is the mean of 10 values acquired at each eye orientation. The torsion values obtained from an uncompensated sampling path are shown in the top row, and both the experimental and simulated data indicate an error magnitude of over 3° at highly eccentric eye positions. Variation in torsion within the 10 measured values at each eye orientation was small, with the mean of the SDs at each eye position being 0.05°. The simulation does not take the illumination of the eye into account, and the close match with the experimental data indicates that the error in torsion was primarily due to incorrect placement of the sampling path due to geometric distortion.

Ocular torsion values derived from a compensated sampling path are shown in the bottom row of Fig. 9. The experimental data demonstrate that the geometric compensation algorithm significantly reduced the torsion error, with an error magnitude of 0.11 ± 0.09°

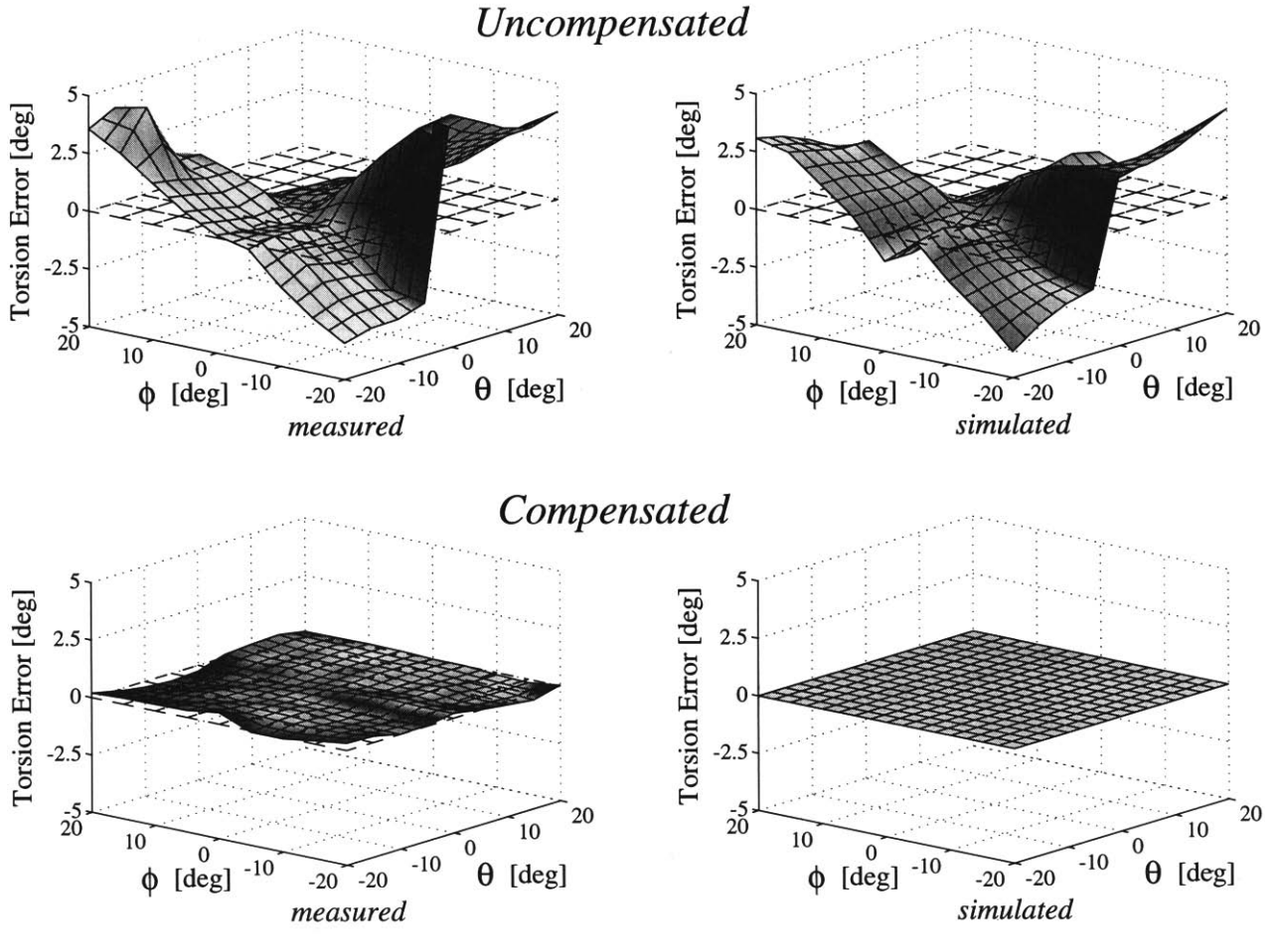

FIGURE 9. *In vitro* validation of the geometric compensation algorithm. An artificial eye was placed in various eccentric positions with zero Fick torsion, and at each position the torsional component of eye position was determined with the VTM system using a geometrically uncompensated sampling path (upper left) and a compensated sampling path (lower left). Each torsion value is the mean of 10 images. The right column shows the results of a numerical simulation of the polar cross correlation method using the same uncompensated (upper right) and compensated (lower right) sampling paths as the VTM system.

(mean and SD, calculated from all 810 torsion values), with a maximum error of 0.53°. Most of the variation in measured torsion occurred at highly eccentric eye positions, with an error magnitude of $0.09 \pm 0.06°$ for eye positions corresponding to θ and ϕ up to $\pm 15°$, increasing to $0.16 \pm 0.13°$ for magnitudes of θ and ϕ between 15° and 20°. As expected, the simulation correctly indicated 0° Fick torsion over the entire range of horizontal and vertical eye positions.

Dual quadrant torsion measurements

To test the geometric compensation algorithm *in vivo*, ocular torsion was measured with the VTM system using sampling paths from both the upper and lower right quadrants simultaneously (see Fig. 6), over a number of horizontal and vertical eye positions. In both of these quadrants iral signatures were obtained from a compensated and uncompensated sampling path. At all eye orientations the Fick torsion determined from sampling paths from different segments of the iris should be identical.

The subject was seated with head and body restrained, and testing carried out in darkness except for an LED array positioned approx. 60 cm in front of the subject's left eye. Three fixation LEDs were positioned in such a manner as to elicit eye positions of approx. +10° horizontal, and between ±10° vertical. Following calibration of the VTM system the subject fixated on various points as directed by the system operator, and at each position 20 images were acquired of the subject's left eye. From each image horizontal, vertical and four torsional Fick angles (from compensated and uncompensated sampling paths from the upper and lower right quadrants) were determined. Accurate fixation was not required for this experiment, since the aim was to obtain *simultaneous* torsion measures from compensated and uncompensated sampling paths from two iral segments.

The results of this experiment for two subjects are shown in Fig. 10. The bar plots in the left column show the difference (mean and SD) between the torsion measured from sampling paths from the upper and lower right quadrants at six different eye orientations. The dark and light bars indicate torsion differences between geometrically compensated and uncompensated sampling paths respectively. The dashed lines show the resolution of the VTM system for the two test runs (0.19°

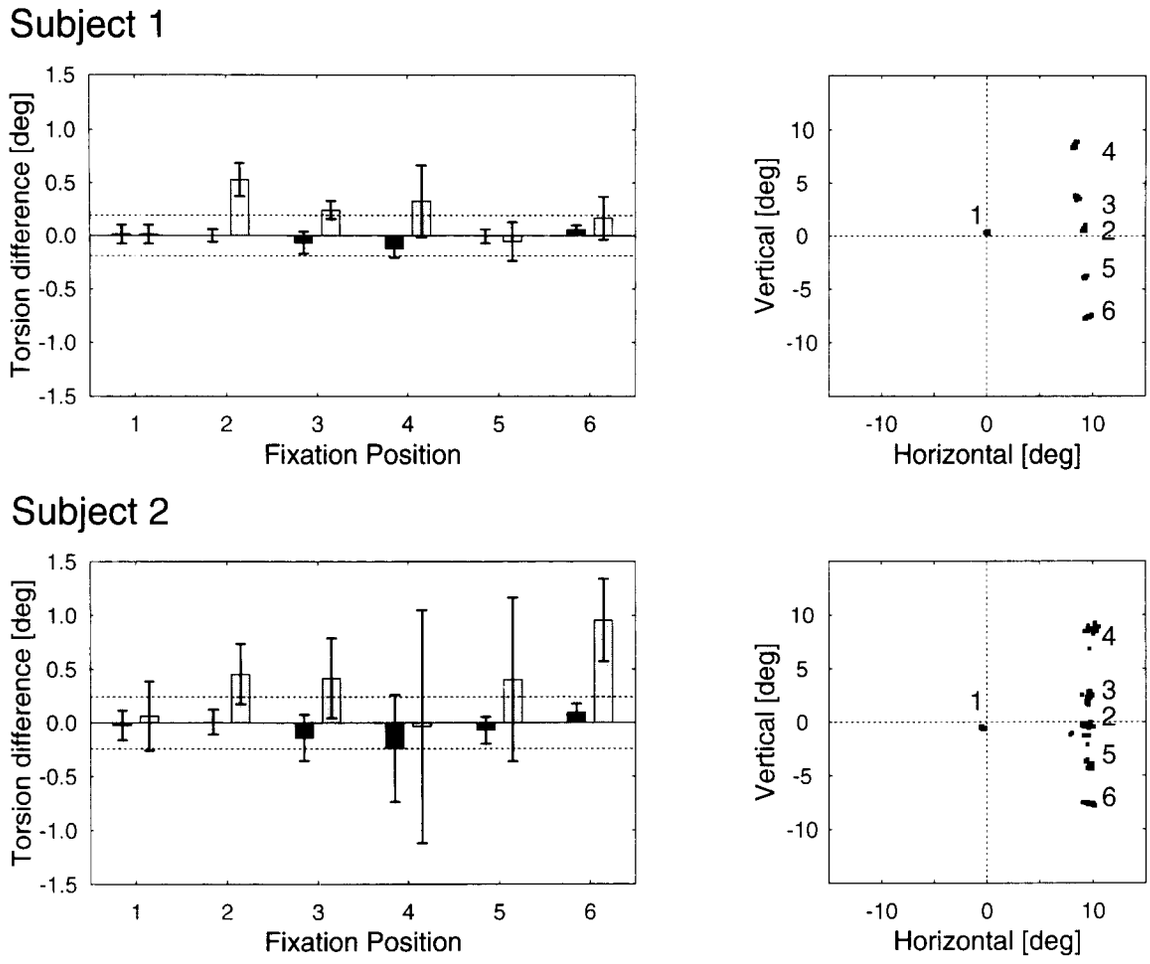

FIGURE 10. *In vivo* validation of the geometric compensation algorithm for two subjects. The plots in the left column show the difference between Fick torsion measured by two geometrically compensated (■) and two uncompensated (□) sampling paths. Each bar indicates the mean and SD of the difference in torsion measured simultaneously from sampling paths in the upper and lower right quadrants (20 values). The dashed lines indicate the resolution of the VTM system for each test. The plots in the right column show the horizontal and vertical Fick coordinates of eye position at each fixation position.

for subject 1 and 0.24° for subject 2). The radial distance r_{sp} from the pupil centre to the sampling paths (and therefore the VTM system resolution) was limited to maintain all four sampling paths within the image plane at all eye orientations. The plots in the right column show the corresponding horizontal and vertical eye positions.

For both subjects the differences between torsion indicated by the two compensated sampling paths were negligible (i.e. within the resolution of the VTM system) at all measured eye orientations. The only exception was the value at position 4 for subject 2, where iral signatures from the upper quadrant were contaminated by blink artifact. For all eccentric eye positions the torsion differences for the two uncompensated sampling paths were significantly larger than for the compensated sampling paths and greater than the VTM resolution.

In summary, the Fick torsion indicated by the compensated sampling paths from the upper and lower right quadrants are in agreement over the six eye orientations, demonstrating the effectiveness of the geometric compensation algorithm *in vivo*. The uncompensated sampling paths indicate *different* torsion values from the two quadrants at eccentric eye positions. This discrepancy is due to geometric distortion of the sampling paths, which is dependent on eye orientation as well as the angular positioning of the sampling arc on the iris.

DISCUSSION

Polar cross correlation is a powerful method for the measurement of ocular torsion, but until now has been restricted to small horizontal and vertical eye movements. We have extended this technique to enable accurate measurement of three-dimensional eye position over a range of eye movements of ±20°, by considering the geometry of the eye when determining the shape and location of the sampling path in the image plane. If eye geometry is not taken into account, large torsional errors of up to 3° arise at eccentric eye positions. Given the small size of ocular torsion (usually in the order of a few degrees), such errors are unacceptable.

Validation of any eye measurement system is an important step, and we have tested our video system VTM both *in vitro* (using an artificial eye) and *in vivo*.

For purely torsional eye movements, we have previously demonstrated that a video system based on polar cross correlation can achieve an accuracy of 0.1° (Moore *et al.*, 1991). The results presented in this paper show that by compensating for the geometric distortion of the sampling path, polar cross correlation can achieve a similar accuracy for torsion measurements over a horizontal and vertical range of eye movements of ±15°. If this accuracy is to be maintained at more eccentric eye positions, the full effects of the rotational offset of the head-fixed coordinate system with respect to the camera have to be taken into consideration. Furthermore, the more accurate central projection must be used in place of the orthographic projection employed in our implementation. Our results with an artificial eye also demonstrate the effectiveness of the centre of mass algorithm for determining the horizontal and vertical components of eye position.

Although measurements with an artificial eye are a useful means of establishing the accuracy of video-based eye position measurement, it is important to include calibration and comparison studies on human eyes. There are a number of important factors which cannot be simulated using an artificial eye, such as drifts and oscillations in torsional eye position (Collewijn *et al.*, 1985; Ferman, Collewijn, & Van den Berg, 1987a; Ferman, Collewijn, Jansen & Van den Berg, 1987b; Markham & Diamond, 1992), infrared reflections on the surface of the eye, and fluctuations in pupil size. We have simultaneously measured eye position using the VTM system and scleral search coils, which have been the accepted standard for three-dimensional eye position measurement, and have found the two techniques to be in close agreement. This result is in accordance with a previous direct comparison of the VTM system with the 35 mm photographic technique for purely torsional eye movements (Diamond, Markham, Simpson & Curthoys, 1979), which showed that the two systems led to almost identical results (Curthoys, Moore, McCoy, Halmagyi, Markham, Diamond, Wade & Smith, 1992).

The methods described above determine the horizontal and vertical orientation of the optical axis of the eye, and ocular torsion about this axis. For applications in which it is important to know exactly where the subject is looking, one has to allow for the fact that the visual axis is orientated nasally with respect to the optical axis, and the angle between these two axes is approx. 5° (Bennett & Rabbetts, 1984).

An unresolved problem is that the eye does not exhibit ideal "ball and socket" behaviour. During vergence, lateral translations of the eye in the orbit of up to 200 μm occur (Enright, 1984), which would appear to a video-based system as a horizontal rotation of around 1°. With any video eye measurement system it is difficult to distinguish between a translation of the rotational centre of the eye, and a rotation about this centre. We have simulated the effect of ocular translation on torsion measurement using polar cross correlation, and found that errors of up to 0.23° are induced (Haslwanter & Moore, 1995). Translations could in practice be compensated for by the combined tracking of the centre of the pupil and of the reflection of an infrared LED on the cornea. If the fixation distance is constant, as was the case in the experiments reported in this paper, this effect is negligible.

While video-based systems still have unresolved problems such as dealing with translations of the eye and low sampling rates, search coils also have their weaknesses, such as difficulties with detecting coil slippage, the expense of the contact lenses, and their quasi-invasive nature. As the development of image processing hardware continues to advance at an increasing rate, the temporal and spatial resolution of video-based systems will constantly improve, and may soon rival the scleral coils technique.

REFERENCES

Bennett, A. G. & Rabbetts, R. B. (1984). *Clinical visual optics*. London: Butterworths.

Clarke, A. H., Teiwes, W. & Scherer, H. (1991). Videooculography—An alternative method for measurement of three-dimensional eye movements. In Schmid, R. & Zambarbieri, D. (Eds), *Oculomotor control and cognitive processes* (pp. 431–443). Amsterdam: Elsevier.

Collewijn, H., Van der Steen, J., Ferman, L. & Jansen, T. C. (1985). Human ocular counterroll: Assessment of static and dynamic properties from electromagnetic scleral coil recordings. *Experimental Brain Research, 59*, 185–196.

Curthoys, I. S., Moore, S. T., McCoy, S. G., Halmagyi, G. M., Markham, C. H., Diamond, S. G., Wade, S. W. & Smith, S. T. (1992). VTM—A new method of measuring ocular torsion using image processing techniques. *Annals of the New York Academy of Sciences, 656*, 826–828.

Diamond, S. G., Markham, C. H., Simpson, N. E. & Curthoys, I. S. (1979). Binocular counterrolling in humans during dynamic rotation. *Acta Otolaryngology, 87*, 490–498.

Enright, J. T. (1984). Saccadic anomalies: Vergence induces large departures from ball-and-socket behaviour. *Vision Research, 24*, 301–308.

Ferman, L., Collewijn, H. & Van den Berg, A. V. (1987a). A direct test of Listing's law—I. Human ocular torsion measured in static tertiary positions. *Vision Research, 27*, 929–938.

Ferman, L., Collewijn, H., Jansen, T. C. & Van den Berg, A. V. (1987b). Human gaze stability in the horizontal, vertical and torsional direction during voluntary head movements, evaluated with a three-dimensional scleral induction coil technique. *Vision Research, 27*, 811–828.

Fick, A. (1854). Die Bewegungen des menschlichen Augapfels. *Zeitschrift für rationelle Medizin, 4*, 109–128.

Gonzalez, R. C. & Wintz, P. (1987). *Digital image processing* (2nd edn). Reading, Mass.: Addison-Wesley.

Gullstrand, A. (1909). Appendix II.3. The optical system of the eye. In von Helmholtz. H., *Handbuch der Physiologischen Optik*. English translation: Southall, J. P. C. (Ed.), *Treatise on physiological optics* (Vol. 1, pp. 350–358). New York: Optical Society of America. Reprinted 1962. New York: Dover.

Haslwanter, T. (1995). Mathematics of 3-dimensional eye rotations. *Vision Research, 35*, 1727–1740.

Haslwanter, T. & Moore, S. T. (1995). A theoretical analysis of three dimensional eye position measurement using polar cross correlation. *IEEE Transactions on Biomedical Engineering*. In press.

Hatamian, M. & Anderson, D. J. (1983). Design considerations for a real-time ocular counterroll instrument. *IEEE Transactions on Biomedical Engineering, BME-30*, 278–288.

Haustein, W. (1989). Considerations on Listing's Law and the primary position by means of a matrix description of eye position control. *Biological Cybernetics, 60*, 411–420.

von Helmholtz, H. (1867). *Handbuch der Physiologischen Optik*.

Leipzig: Voss. English translation: Southall, J. P. C. (Ed.), *Treatise on physiological optics*. New York: Optical Society of America. Reprinted 1962, New York: Dover.

Markham, C. H. & Diamond, S. G. (1992). Further evidence to support disconjugate eye torsion as a predictor of space motion sickness. *Aviation Space and Environmental Medicine, 63*, 118–121.

Merfeld, D. M. & Young, L. R. (1992). Three dimensional eye velocity measurements following postrotational tilt in the monkey. *Annals of the New York Academy of Sciences, 656*, 783–794.

Moore, S. T., Curthoys, I. S. & McCoy, S. G. (1991). VTM—An image-processing system for measuring ocular torsion. *Computer Methods and Programs in Biomedicine, 35*, 219–230.

Nakayama, K. (1974). Photographic determination of the rotational state of the eye using matrices. *American Journal of Optometry and Physiological Optics, 51*, 736–742.

Ott, D., Gehle, F. & Eckmiller, R. (1990). Video-oculographic measurement of 3-dimensional eye rotations. *Journal of Neuroscience Methods, 35*, 229–234.

Parker, J. A., Kenyon, R. V. & Young, L. R. (1985). Measurement of torsion from multitemporal images of the eye using digital signal processing techniques. *IEEE Transactions on Biomedical Engineering, BME 32*, 28–35.

Remmel, R. S. (1984). An inexpensive eye movement monitor using the scleral search coil technique. *IEEE Transactions on Biomedical Engineering, BME-31*, 388–390.

Robinson, D. A. (1963). A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Transactions on Biomedical Electronics, BME-10*, 137–145.

Sung, K. & Anderson, D. J. (1991). Analysis of two video eye tracking algorithms. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (Vol. 13, No. 5, pp. 1949–1950).

Vieville, T. & Masse, D. (1987). Ocular counter-rolling during active head tilting in humans. *Acta Otolaryngology, 103*, 280–290.

Westheimer, G. (1957). Kinematics of the eye. *Journal of the Optical Society of America, 47*, 967–974.

Wolff, E. (1940). *The anatomy of the eye and orbit* (2nd edn). London: H. K. Lewis & Co.

Yamanobe, S., Taira, S., Morizono, T., Yagi, T. & Kamio, T. (1990). Eye movement analysis system using computerized image recognition. *Archives of Otolaryngology—Head and Neck Surgery, 116*, 338–341.

APPENDIX A

Rotation Convention

To uniquely characterize the three-dimensional position of the eye, we introduce an eye-fixed coordinate system $\{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$, which is oriented such that $\underline{e}_1$ passes through the centre of the pupil, and that $\{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$ coincides with the head-fixed coordinate system $\{\underline{h}_1, \underline{h}_2, \underline{h}_3\}$ when the eye is in the reference position. The pupil centre in head-fixed coordinates in this position is $\underline{P}_0 = (r_p, 0, 0)$, where r_p is the distance between the pupil centre and the centre of the eye. Three-dimensional eye position can be described by a rotation matrix $\mathbb{R}$ which characterizes the rotation from the reference position $\underline{P}_0$ to the current point $\underline{P}$

$$\underline{P} = \mathbb{R} \cdot \underline{P}_0. \tag{A1}$$

Due to the non-commutativity of finite rotations, it is necessary to define a rotation convention. We have chosen the *Fick sequence* (Fick, 1854), which is often used in oculomotor research, where $\mathbb{R}$ is composed of a horizontal rotation about $\underline{e}_3$ by θ,

$$\mathbb{R}(\underline{e}_3, \theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{A2}$$

followed by a vertical rotation about the rotated axis $\underline{e}_2$ by ϕ,

$$\mathbb{R}(\underline{e}_2, \phi) = \begin{bmatrix} \cos(\phi) & 0 & \sin(\phi) \\ 0 & 1 & 0 \\ -\sin(\phi) & 0 & \cos(\phi) \end{bmatrix} \tag{A3}$$

and a torsional rotation about the twice rotated optical axis $\underline{e}_1$ by ψ.

$$\mathbb{R}(\underline{e}_1, \psi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\psi) & -\sin(\psi) \\ 0 & \sin(\psi) & \cos(\psi) \end{bmatrix}. \tag{A4}$$

Such descriptions of rotations, in which consecutive rotations are executed about eye-fixed axes, are often called *passive rotations* or *rotations of the coordinate system*. According to the right-hand rule, eye movements to the left, down and clockwise (from the subject's point of view) are positive. Combined horizontal and vertical eye movements are described by multiplication of the rotation matrices of equations (A2) and (A3)

$$\mathbb{R}_{\theta\phi} = \mathbb{R}(\underline{e}_3, \theta) \cdot \mathbb{R}(\underline{e}_2, \phi) = \begin{bmatrix} \cos(\theta)\cos(\phi) & -\sin(\theta) & \cos(\theta)\sin(\phi) \\ \sin(\theta)\cos(\phi) & \cos(\theta) & \sin(\theta)\sin(\phi) \\ -\sin(\phi) & 0 & \cos(\phi) \end{bmatrix} \tag{A5}$$

and three-dimensional eye position is characterized by the Fick matrix

$$\mathbb{R}_{\theta\phi\psi} = \mathbb{R}(\underline{e}_3, \theta) \cdot \mathbb{R}(\underline{e}_2, \phi) \cdot \mathbb{R}(\underline{e}_1, \psi) = \begin{bmatrix} \cos(\theta)\cos(\phi) & \cos(\theta)\sin(\phi)\sin(\psi) - \sin(\theta)\cos(\psi) & \cos(\theta)\sin(\phi)\cos(\psi) + \sin(\theta)\sin(\psi) \\ \sin(\theta)\cos(\phi) & \sin(\theta)\sin(\phi)\sin(\psi) + \cos(\theta)\cos(\psi) & \sin(\theta)\sin(\phi)\cos(\psi) - \cos(\theta)\sin(\psi) \\ -\sin(\phi) & \cos(\phi)\sin(\psi) & \cos(\phi)\cos(\psi) \end{bmatrix}. \tag{A6}$$

Acknowledgements—The authors gratefully acknowledge the help of many people and organizations: B. Denby (Department of Meterology, Utrecht University) for his valuable advice on camera optics; G. M. Halmagyi, Mark Hedley, M. J. Todd and L. McGarvie (the University of Sydney and the Royal Prince Alfred Hospital). This work is supported by the NH and MRC of Australia and the Australian Brain Foundation.

APPENDIX B

Central vs Orthographic Projection

Nakayama (1974) has suggested that the simpler orthographic projection can be used in the determination of eye position instead of the central projection of equation (3). We have investigated this hypothesis by determining the error in the coordinates of the pupil centre in the image plane if an orthographic projection is used. With the imaging geometry of Fig. 1, and the centre of the eye translated

Equation (A6) is also used to describe the rotational offset $\mathbb{R}_{off}$ of $\{\underline{h}_1, \underline{h}_2, \underline{h}_3\}$ with respect to the camera-fixed coordinate system $\{\underline{c}_1, \underline{c}_2, \underline{c}_3\}$. For an extensive description of the mathematical representation of eye movements see Haslwanter (1995).

by $\underline{T} = (-(f+d), t_2, t_3)$ with respect to the camera-fixed coordinate frame, the central projection of the pupil centre $\underline{P}$ onto the image plane at a particular eye orientation (θ, ϕ) is given by

$$\begin{pmatrix} 0 \\ x_{cent} \\ y_{cent} \end{pmatrix} = \left(\frac{f}{d - r_p}\right) * \frac{1}{1 + E(\theta, \phi, d, r_p)} * \begin{pmatrix} 0 \\ -(p_2 + t_2) \\ -(p_3 + t_3) \end{pmatrix} \tag{B1}$$

where $E(\theta, \phi, d, r_p) = (r_p/(d - r_p)) * [1 - \cos(\theta) * \cos(\phi)]$.
The corresponding orthographic projection of $\underline{P}$ is

$$\begin{pmatrix} 0 \\ x_{orth} \\ y_{orth} \end{pmatrix} = k * \begin{pmatrix} 0 \\ p_2 + t_2 \\ p_3 + t_3 \end{pmatrix} \tag{B2}$$

where k is a scaling factor related to the image magnification. If k is chosen to be $f/(d - r_p)$, the error in the pupil centre coordinates in the image plane using an orthographic projection (relative to the values given by the more accurate central projection) is given by E, which is positive over the natural range of eye movements.

Using $r_p = 12$ mm for the radius of the eye at the pupil centre (Wolff, 1940), and a typical distance between the lens plane and the centre of the eye for the VTM system of $d = 72$ mm, the maximum error induced by assuming an orthographic projection is 0.6% for eye positions corresponding to θ and ϕ up to $\pm 10^\circ$, 1.3% for $\pm 15^\circ$ and 2.3% for $\pm 20^\circ$. These errors can be reduced further by increasing d.

APPENDIX C

Determination of $\mathbb{R}_{off}$ and $\underline{T}$

Based on the results of Appendix B, the central projection of equation (3) can be replaced by the simpler orthographic projection of equation (B2) if the distance d between the lens plane and the centre of the eye is much larger than the radius of the eye r_p, or for small eye movements (θ and ϕ up to $\pm 10^\circ$). Equations (3) and (4), which are necessary for calculating the pupil centre in the image plane $\underline{P}''$ at a particular known eye orientation (θ, ϕ), are then reduced to

$$\underline{P}'' = \begin{pmatrix} 0 \\ x \\ y \end{pmatrix} = r_p * \begin{pmatrix} 0 \\ R_{21} \\ R_{31} \end{pmatrix} + \begin{pmatrix} 0 \\ x_0 \\ y_0 \end{pmatrix} \tag{C1}$$

where $\underline{T}'' = (0, x_0, y_0)$ is the projection of the centre of the eye onto the image plane, and R_{ij} is the matrix element in the ith row and jth column of the rotation matrix $\mathbb{R}$, given by

$$\mathbb{R} = \mathbb{R}_{off} \cdot \mathbb{R}_{\theta\phi}. \tag{C2}$$

We can now derive analytical solutions for the six parameters which need to be determined for the calibration of the system: the Fick angles of the offset matrix $\mathbb{R}_{off}$, $(\theta_{off}, \phi_{off}, \psi_{off})$, the radius of the eye at the pupil centre r_p, and $\underline{T}'' = (0, x_0, y_0)$.

Four calibration positions are needed to determine these six unknown parameters. We have chosen two purely horizontal positions at $\pm\theta_{cal}$, and two purely vertical positions at $\pm\phi_{cal}$, where $\theta_{cal} = \phi_{cal}$. This gives us five points: four corresponding to the location of the pupil centre in the image plane at each of the four calibration angles $(0, x_{+\theta}, y_{+\theta})$, $(0, x_{-\theta}, y_{-\theta})$, $(0, x_{+\phi}, y_{+\phi})$ and $(0, x_{-\phi}, y_{-\phi})$, plus the location of the pupil centre with the eye in the reference position $(0, x_r, y_r)$. From this set of points and equations (C1) and (C2) the required parameters are derived as:

$$\psi_{off} = -\text{atan}\left(\frac{y_{+\theta} - y_{-\theta}}{y_{+\phi} - y_{-\phi}}\right) \tag{C3}$$

$$r_p^2 = \left(\frac{y_{+\phi} + y_{-\phi} - 2 * y_r}{2 * (1 - \cos(\phi_{cal}))}\right)^2 + \left(\frac{y_{+\phi} - y_{-\phi}}{2 * \cos(\psi_{off}) * \sin(\phi_{cal})}\right)^2 \tag{C4}$$

$$\phi_{off} = \text{asin}\left(\frac{y_{+\phi} + y_{-\phi} - 2 * y_r}{2 * r_p * (1 - \cos(\phi_{cal}))}\right) \tag{C5}$$

$$\theta_{off} = \text{asin}\left(\frac{x_{+\phi} + x_{-\phi} - 2 * x_r}{2 * r_p * \cos(\phi_{off}) * (\cos(\phi_{cal}) - 1)}\right) \tag{C6}$$

$$\underline{T}'' = \begin{pmatrix} 0 \\ x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 0 \\ x_r \\ y_r \end{pmatrix} - r_p * \begin{pmatrix} 0 \\ \cos(\phi_{off}) * \sin(\theta_{off}) \\ -\sin(\phi_{off}) \end{pmatrix}. \tag{C7}$$

If the central projection of equation (3) is to be used in subsequent calculations, the component of the translation vector $\underline{T}$ along the $\underline{c}_1$ axis, $t_1 = -(f + d)$, must be determined by measuring the distance d between the lens plane and the centre of the eye.

APPENDIX D

Formulae for Determining Eye Position

In the practical implementation of a real-time video system for measuring eye position it is important to minimize the amount of processing in the data acquisition phase. In our video system (VTM) the computationally expensive tasks of calculating the cross correlation of two iral signatures, and compensating for the rotational offset $\mathbb{R}_{off}$ of the head-fixed coordinate system with respect to the camera-fixed frame, are carried out after testing. To do this, we set $\mathbb{R}_{off} = \mathbb{I}$ (i.e. the unit matrix) during data acquisition. This has the effect that eye positions are not expressed with respect to the stereotaxically defined head-fixed coordinate system $\{\underline{h}_1, \underline{h}_2, \underline{h}_3\}$, but with respect to a rotated coordinate system $\{\underline{h}_1^r, \underline{h}_2^r, \underline{h}_3^r\}$, which is translated by $\underline{T} = (-(f + d), t_2, t_3)$ with respect to the camera-fixed coordinate system and oriented such that $\{\underline{h}_1^r, \underline{h}_2^r, \underline{h}_3^r\}$ are parallel to $\{\underline{c}_1, \underline{c}_2, \underline{c}_3\}$. The Fick angles of eye position in this coordinate frame are denoted $\{\theta^r, \phi^r, \psi^r\}$, and the relationship to $\{\theta, \phi, \psi\}$, the Fick angles with respect to $\{\underline{h}_1, \underline{h}_2, \underline{h}_3\}$, is given by

$$\mathbb{R}_{\theta\phi\psi} = \mathbb{R}_{off}^{-1} \cdot \mathbb{R}_{\theta\phi\psi}^r \tag{D1}$$

where $\mathbb{R}_{\theta\phi\psi}^r$ is the Fick matrix obtained by inserting $\{\theta^r, \phi^r, \psi^r\}$ into equation (A6). A practical implementation of a geometric approach to three-dimensional eye position measurement proceeds as follows.

(i) Iral reference signature. When the eye is in the reference position, the horizontal and vertical Fick angles with respect to $\{\underline{h}_1^r, \underline{h}_2^r, \underline{h}_3^r\}$ are given by $(\theta_{off}, \phi_{off})$, determined during calibration of the system. A sampling arc $\underline{S}_0$ is formed in $\{\underline{h}_1^r, \underline{h}_2^r, \underline{h}_3^r\}$ using equation (5), and the coordinates of the sampling arc $\underline{S}'$ in the camera-fixed coordinate frame can be determined using equation (7) as

$$\underline{S}' = \begin{pmatrix} r_p * \cos(\theta^r) * \cos(\phi^r) - r_{ip} * \cos(\alpha) * \sin(\theta^r) + r_{ip} * \sin(\alpha) * \cos(\theta^r) * \sin(\phi^r) + t_1 \\ r_p * \sin(\theta^r) * \cos(\phi^r) + r_{ip} * \cos(\alpha) * \cos(\theta^r) + r_{ip} * \sin(\alpha) * \sin(\theta^r) * \sin(\phi^r) + t_2 \\ r_{ip} * \sin(\alpha) * \cos(\phi^r) - r_p * \sin(\phi^r) + t_3 \end{pmatrix} \tag{D2}$$

where $\theta^r = \theta_{off}$ and $\phi^r = \phi_{off}$. The iral reference signature $I_{ref}(\alpha)$ is obtained from the sampling path $\underline{S}''$, the projection of $\underline{S}'$ onto the image plane [given by equation (3)], and stored for later processing.

(ii) Horizontal and vertical eye position. For subsequent video images it is first necessary to determine the horizontal and vertical Fick angles (θ^r, ϕ^r) at the current eye position. After an image is acquired the projection of the centre of the pupil in the image plane $\underline{P}'' = (0, x, y)$ is determined using the centre of mass algorithm. If β is defined as the angle between the optical axis and the $\underline{h}_1^r$ axis, and noting that the $\underline{h}_1^r$ component of the pupil centre $\underline{P}$ with respect to $\{\underline{h}_1^r, \underline{h}_2^r, \underline{h}_3^r\}$ is given by $p_1 = r_p * \cos(\beta)$, $\underline{P}$ can be obtained from equations (2) and (3) as

$$\underline{P} = \begin{pmatrix} r_p * \cos(\beta) \\ \left(\frac{1}{f}\right) * (x * r_p * \cos(\beta) - (x - x_0) * d) \\ \left(\frac{1}{f}\right) * (y * r_p * \cos(\beta) - (y - y_0) * d) \end{pmatrix} \tag{D3}$$

where $\underline{T}'' = (0, x_0, y_0)$ is the projection of the centre of the eye onto the image plane. Using the fact that $\sqrt{p_2^2 + p_3^2} = r_p * \sin(\beta)$, $\cos(\beta)$ can be derived as

$$\cos(\beta) = \frac{d}{r_p * (x^2 + y^2 + f^2)} * \left(\xi + \sqrt{\xi^2 - (x^2 + y^2 + f^2) * \left(\delta_x^2 + \delta_y^2 - \left(\frac{f^2 * r_p^2}{d^2} \right) \right)} \right) \quad \text{(D4)}$$

where $\delta_x = x - x_0$, $\delta_y = y - y_0$ and $\xi = x * \delta_x + y * \delta_y$. By substituting equation (D3) into equation (6) the horizontal and vertical Fick angles (θ^r, ϕ^r) of the current eye position are given by

$$\begin{pmatrix} \theta^r \\ \phi^r \end{pmatrix} = \begin{bmatrix} \operatorname{asin}\left(\frac{p_2}{r_p * \cos(\phi^r)} \right) \\ -\operatorname{asin}\left(\frac{p_3}{r_p} \right) \end{bmatrix}. \quad \text{(D5)}$$

Note that if an orthographic projection is used, $\underline{P}$ can be obtained directly from $\underline{P}''$ by inverting equation (B2).

(iii) Iral signature. The location of the sampling arc $\underline{S}'$ in camera-fixed coordinates is calculated by substituting (θ^r, ϕ^r) into equation (D2). The geometrically compensated sampling path $\underline{S}''$ is given by applying the central projection of equation (3) to $\underline{S}'$, and the iral signature $I(\alpha)$ is obtained by sampling pixel intensity along $\underline{S}''$.

(iv) Offline processing. During data acquisition, the horizontal and vertical Fick angles (θ^r, ϕ^r) and the iral signature $I(\alpha)$ are stored for each video frame. After testing, the Fick torsion ψ^r of eye position with respect to $\{\underline{h}_1^r, \underline{h}_2^r, \underline{h}_3^r\}$ is determined as $\psi^r = \psi_{\text{off}} + \psi_{cc}$, where ψ_{cc} is the shift in the peak of the cross correlation of each iral signature $I(\alpha)$ with the iral reference signature $I_{ref}(\alpha)$. Using equation (A6), the Fick angles (θ, ϕ, ψ) of eye position with respect to $\{\underline{h}_1, \underline{h}_2, \underline{h}_3\}$ can be derived as

$$\begin{pmatrix} \theta \\ \phi \\ \psi \end{pmatrix} = \begin{bmatrix} \operatorname{atan}\left(\frac{R_{21}}{R_{11}} \right) \\ -\operatorname{asin}(R_{31}) \\ \operatorname{atan}\left(\frac{R_{32}}{R_{33}} \right) \end{bmatrix} \quad \text{(D6)}$$

where $\mathbb{R}$ is the matrix $\mathbb{R}_{\theta\phi\psi}$ given by equation (D1).